APPLICATION FOR
UNITED STATES PATENT
IN THE NAME

Of

ROBERT NOVAK

For

SYSTEM AND METHOD FOR A SOFTWARE STEERABLE WEB CAMERA

DOCKET NO. 52126.00002

Please direct communications to:
SQUIRE, SANDERS & DEMPSEY L.L.P.
600 Hansen Way
Palo Alto, CA 94304-1043
(650) 856-6500

Express Mail Number: EL701362645US

SYSTEM AND METHOD FOR A SOFTWARE STEERABLE WEB CAMERA

Inventor: Robert Novak

Technical Field

This disclosure relates generally to digital imaging, digital video or web cameras, and more particularly but not exclusively, to systems and methods for capturing camera images by use of software control.

Background

Conventional digital imaging, digital video or web cameras ("webcams") can be used for teleconferencing, surveillance, and other purposes. One of the problems with conventional webcams is that they have a very restricted field of vision. This restricted vision field is due to the limitations in the mechanism used to control the webcam and in the optics and other components in the webcam.

In order to increase the vision field of a webcam, the user might manually control the webcam to pan and/or tilt in various directions (e.g., side-to-side or up-and-down) and/or to zoom in or away from an image to be captured. However, this manual technique is inconvenient, as it requires the user to stop whatever he/she is doing, to

readjust the webcam, and to then resume his/her previous activity.

Various other schemes have been proposed to increase the webcam vision field, such as adding complex lens assemblies and stepper motors to the webcams to permit the camera to perform the pan and zoom functions. However, complex lens assemblies are expensive and will make webcams unaffordable for many consumers. Additionally, stepper motors use moving or mechanical parts that may fail after a certain amount of time, thus requiring expensive repairs or the need to purchase a new webcam. Stepper motors may also disadvantageously suffer from hysterisis, in which repeated pan, tilt or zooming operations lead to slightly inconsistent settings during each operation.

Furthermore, repairs for webcams on set top boxes (STBs) are particularly expensive because of the required service call for repairing the STB webcam.

Accordingly, there is need for a new system and method to allow webcams to increase their vision field. There is also a need for a new system and method to permit webcams to perform particular operations, such as panning, tilting, and/or zooming, without using stepper motors or requiring the user to physically adjust the webcam.

BRIEF DESCRIPTION OF THE DRAWINGS

Non-limiting and non-exhaustive embodiments of the present invention are described with reference to the following figures, wherein like reference numerals refer to like parts throughout the various views unless otherwise specified.

Figure 1 is a block diagram showing a webcam coupled to a set top box according to an embodiment of the invention.

Figure 2 is a block diagram of an embodiment of the webcam of Figure 1.

Figure 3 is a block diagram of an embodiment of the set top box of Figure 1.

Figure 4 is a block diagram of one example of a memory device of the set top box.

Figure 5A is an illustrative example block diagram showing a function of the webcam of Figure 1 in response to particular pan and/or tilt commands.

Figure 5B is an illustrative example block diagram of selected subsets in a digitized scene image data in response to particular pan and/or tilt commands.

Figure 6A is an illustrative example block diagram of a selected subset image data with distortions.

Figure 6B is an illustrative example block diagram of a selected subset image data that has been distortion compensated.

Figure 7 is a flowchart diagram of a method according to an embodiment of the invention.

Figure 8A is an illustrative example block diagram showing a function of the webcam of Figure 1 in response to particular pan and zoom commands.

Figure 8B is an illustrative example block diagram of a selected subset in the digitized scene image data in response to a particular pan command;

Figure 8C is an illustrative example block diagram of the selected subset in Figure 8B in response to a particular zoom command.

Figure 9 is an illustrative example block diagram of the selected subset in Figure 9 in response to another particular zoom command.

Figure 10 is a flowchart diagram of a method according to another embodiment of the invention.

Figure 11 is another diagram shown to further assist in describing an operation of an embodiment of the invention.

DETAILED DESCRIPTION OF THE ILLUSTRATED EMBODIMENTS

Embodiments of a system and method for a software steerable camera are disclosed herein. As an overview, an embodiment of the invention provides a system and method that capture camera images by use of software control. As an example, the camera may be web camera or other types of camera that can support a wide angle lens. The wide angle lens is used to capture a scene or image in the wide field of vision. The captured scene or image data is then stored in an image collection array and then digitized and stored in memory. In one embodiment, the image collection array is a relatively larger sized array to permit the array to store image data from the wide vision field. Processing is performed for user commands to effectively pan the webcam in particular directions and/or to zoom the webcam toward or away from an object to be captured as an image. However, instead of physically moving the webcam in response to the user commands, a particular subset of the digitized data is selected and processed so that selected subset data provides a simulated panning and/or zooming of the image of the captured object. A compression/correction engine can then compensate the selected subset data for distortion and compress the selected subset data for transmission.

The invention advantageously permits a camera, such as a webcam, to have a wide vision field. The invention may also advantageously provide a wide vision field for cameras that have short depth fields. The invention also advantageously avoids the use of stepper motors to obtain particular images based on pan and zoom commands from the user.

In the description herein, numerous specific details are provided, such as the description of system components in Figures 1 through 10, to provide a thorough understanding of embodiments of the invention. One skilled in the relevant art will recognize, however, that the invention can be practiced without one or more of the specific details, or with other methods, components, materials, parts, and the like. In other instances, well-known structures, materials, or operations are not shown or described in detail to avoid obscuring aspects of the invention.

Reference throughout this specification to "one embodiment" or "an embodiment" means that a particular feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment of the present invention. Thus, the appearances of the phrases "in one embodiment" or "in an embodiment" in

various places throughout this specification are not necessarily all referring to the same embodiment. Furthermore, the particular features, structures, or characteristics may be combined in any suitable manner in one or more embodiments.

Figure 1 is a block diagram showing a webcam 100 communicatively coupled to a set top box ("STB") 140 according to an embodiment of the invention. The webcam 100 can capture an image of an object 130 that is in the webcam field of vision. Webcam 100 is communicatively coupled to STB 140 via, for example, a cable 110. Webcam 100 may also be communicatively coupled to STB 140 by use of other suitable connections or methods, such as IR beams, radio signals, suitable wireless transmission techniques, and the like. Typically, STB 140 is communicatively coupled to a cable network 160 and receives TV broadcasts, as well as other data, from the cable network 160. Typically, STB 140 is also communicatively coupled to the Internet 150 or other networks for sending and receiving data. Data received from the Internet 150 or cable network 160 may be displayed on a display 120. STB 140 may also transmit images that are captured by the webcam 100 to other computers via the Internet 150. STB may also transmit the captured webcam images to a printer 165 and/or

to other devices 170 such as a computer in a local area network.

It is noted that embodiments of the invention may also be implemented in other types of suitable cameras that can support a wide angle lens. For example, an embodiment of the invention may be implemented in, for example, security cameras, ATM cash machine cameras, spy cameras, portable cameras, or pin-hole type cameras. It is further noted that the invention is not limited to the use of STB 140. Other processing device may be used according to embodiments of the invention to perform image distortion compensation, image compression, and/or other functions that will be described below.

Figure 2 is a block diagram of an embodiment of the webcam 100 of Figure 1. Webcam 100 comprises a lens 210; a shutter 220; a filter 230; an image collection array 240; a sample stage 245; and an analog to digital converter ("ADC") 250. The lens 210 may be a wide angle lens, such as a fish-eye lens, that has angular field of, for example, at least about 140 degrees, as indicated by lines 200. Using a wide-angle lens allows webcam 100 to capture a larger image area than a conventional webcam. Shutter 220 opens and closes at a pre-specified rate, allowing light into the interior of webcam 100 and onto a filter 230.

Filter 230 allows for image collection array 240 to capture different colors of an image and may include a static filter, such as a Bayer filter, or may include a spinning disk filter. In another embodiment, the filter may be replaced with a beam splitter or other color differentiation device. In another embodiment, webcam 100 does not include a filter or other color differentiation device.

In one embodiment, the image collection array 240 can include charge coupled device ("CCD") sensors or complementary metal oxide semiconductor ("CMOS") sensors, which are generally much less expensive than CCD sensors but may be more susceptible to noise. Other types of sensors may be used in the image collection array 240. The size of the image collection array 240 is relatively larger in size such as, for example, 1024 by 768, 1200 by 768, or 2000 by 1000 sensors. The large sized array permits the array 240 to capture images in the wide vision field 200 that is viewed by the webcam 200.

A sample stage 245 reads the image data from the image collection array 240 when shutter 220 is closed, and an analog-to-digital converter (ADC) 250 converts the image data from an analog to digital form, and feeds the digitized image data to STB 140 via cable 110 for

processing and/or transmission. In an alternative embodiment, the image data may be processed entirely by components of the webcam 100 and transmitted from webcam 100 to other devices such as the printer 165 or computer 170.

For purposes of explaining the functionality of embodiments of the invention, other conventional components that are included in the webcam 100 have been omitted in the figures and are not discussed herein.

Figure 3 is a block diagram of an embodiment of the set top box (STB) 140. STB 140 includes a network interface 300; a processor 310; a memory device 320; a frame buffer 330; a converter 340; a modem 350; a webcam interface 360, and an input device 365, all interconnected for communication by system bus 370. Network interface 300 connects the STB 140 to the cable network 160 (Figure 1) to receive videocasts from the cable network 160. In alternative embodiments, the modem 350 or converter 340 may provide some or all of the functionality of the network interface 300.

Processor 310 executes instructions stored in memory 320, which will be discussed in further detail in conjunction with Figure 4. Frame buffer 330 holds preprocessed data received from webcam 100 via webcam

interface 360. In another embodiment, the frame buffer 330 is omitted since the data from webcam 100 may be loaded into memory 320 instead of loading the data into the frame buffer 330.

Converter 340 can convert, if necessary, digitally encoded broadcasts to a format usable by display 120 (Figure 1). Modem 350 may be a conventional modem for communicating with the Internet 150 via a publicly switched telephone network. The modem 350 can transmit and receive digital information, such as television scheduling information, the webcam 100 output images, or other information to Internet 150. Alternatively, modem 350 may be a cable modem or a wireless modem for sending and receiving data from the Internet 150 or other network.

Webcam interface 360 is communicatively coupled to webcam 100 and receives image output from the webcam 100. Webcam interface 360 may include, for example, a universal serial bus (USB) port, a parallel port, an infrared (IR) receiver, or other suitable device for receiving data. Input device 365 may include, for example, a keyboard, mouse, joystick, or other device or combination of devices that a user (local or remote) uses to control the pan, tilt, and/or zoom webcam 100 by use of software control according to embodiments of the invention. Alternatively,

input device 365 may include a wireless device, such an infrared IR remote control device that is separate from the STB 140. In this particular alternative embodiment, the STB 140 also may include an IR receiver communicatively coupled to the system bus 370 to receive IR signals from the remote control input device.

The components shown in Figure 3 may be configured in other ways and in addition, the components may also be integrated. Thus, the configuration of the STB 140 in Figure 3 is not intended to be limiting.

Figure 4 is a block diagram of an example of a memory device 320 of the set top box 140. Memory device 320 may be, for example, a hard drive, a disk drive, random access memory ("RAM"), read only memory ("ROM"), flash memory, or any other suitable memory device, or any combination thereof. Memory device 320 stores, for example, a compression/correction engine 400 that performs compression and distortion compensation on the image data received from webcam 100. Memory device 320 also stores, for example, a webcam engine 410 that accepts and process user commands relating to the pan, tilt, and/or zoom functions of the webcam 100, as described below. It is also noted the compression/correction engine 400 and/or the webcam engine 410 may be stored in other storage areas that are

accessible by the processor 310. It is noted that either one of the compression/correction engine 400 or webcam engine 410 may be implemented, for example, as a program, module, instruction, or the like.

Compression/correction engine 400 uses, for example, any known suitable skew correction algorithm that compresses a subset of the image output from webcam 100 and that compensates the subset image output for distortion. The distortion compensation of the subset image output may be performed before the compression of the subset image output. In another embodiment, the distortion is automatically corrected in the subset image output when performing the compression of the subset image output, and this leads to a saving in processor resource.

Webcam engine 410 accepts input from a user including instructions to pan the webcam 100 in particular directions and/or to zoom the webcam 100 toward or away from an object to be captured as an image.

Figures 5A and 5B illustrate examples of operations of embodiments of the invention. For example, Figure 5A is a block diagram illustrating a top view of webcam 100. The vision field 200 of the wide angle lens 210 of webcam 100 captures a wide scene area including the three objects 480, 482, and 484. In contrast, a conventional webcam may only

be able to capture the scene area in the limited vision field 481. As a result, a conventional webcam may need manual adjustment or movement by stepper motors to capture the objects 480 or 484 that are outside the limited vision field 481.

For the webcam 100, the entire scene captured in the vision field 200 is stored as an image in the image collection array 240 (Figure 2) and processed by stages 245 and 250, and the image data of the entire scene is stored as digitized scene image data 485 in frame buffer 330 (or memory 320). Thus, each position in the scene area that is covered by vision field 200 corresponds to a position in the image collection array 240 (Figure 2). The values in the positions in the image collection array 240 are then digitized as values of the digitized scene image data 485.

The webcam engine 410 (Figure 4) allows a user to select a subset area in the vision field 200 for display or transmission, so as to simulate a panning/tilting feature of conventional webcams that use stepper motors. For example, assume that the digitized image data 485 was captured in response to a user directly or remotely sending a command 486 via input device 365 to pan the webcam 100 to the left in order to permit the capture of an image of the object 480. The webcam engine 410 receives the pan left

command 486 and accordingly samples an area 487 that contains an image of the object 480 in the digitized scene image data 485.

As another example, if the user were to send a pan right command 488 to webcam 100, then the webcam engine 410 selects an area (subset) 489 that contains an image of the object 484 in the digitized scene image data 485.

As another example, if the user were to send a tilt down command 495 to webcam 100, then the webcam engine 410 selects a subset 496 that contains an image of the bottom portion 498 of object 484 in the digitized scene image data 485.

Webcam engine 410 then passes a selected area (e.g., selected area 487, 489, 496) to the compression/correction engine 400 (Figure 4). The compression/correction engine 400 then performs compression operation and distortion compensation. For example, in Figure 6A, assume that the selected area 487 shows distortions 490 in the image of 480 as a result of using the wide angle lens 210. For images captured by a wide angle lens, the distortions become more pronounced toward the edges of the images. The compression/correction engine 400 can perform distortion compensation to reverse the distortion caused by the wide angle lens 210 on the captured image of object 480.

Typically, this compensation is performed by changing the curved surface of an image into a straight surface.

Figure 6B shows an image of the object 480 without distortions after applying distortion compensation on the selected area 487. Thus, the image of the object 480 is shown as a normal rectilinear image. The selected area 487 can then be compressed by the compression/correction engine 400. In another embodiment, the compression and distortion compensation for selected area 487 can be performed concurrently. In yet another embodiment, the distortion compensation for selected area 487 can be performed before compression of the selected area 487.

The webcam engine 410 then passes the compressed distortion-compensated selected image data 487 to an output device, such as display 120 (Figure 1) for viewing, or to the printer 165 or other devices such as computer 170. In addition to or instead of passing the compressed distortion-compensated selected image data 487 to an output device, webcam engine 410 may transmit the data 487 to another device coupled to the Internet 150.

Figure 7 is a flowchart diagram of a method 600 to perform a panning, tilting or zooming function according to an embodiment of the invention. A user first sends (605) a pan/tilt command indicating a direction of an object to be

captured in an image by a webcam. A scene in the field of vision of a lens of the webcam is then captured (605). In one embodiment, the captured scene is in the vision field 200 (Figure 2) of a wide angle lens 210 of the webcam 100. The captured scene in the vision field is then stored (615) as scene image data in an image collection array. The image collection array may, for example, include charge coupled devices or complementary metal oxide semiconductor sensors. The scene image data in the image collection array is then processed and stored (620) as a digitized scene image data. The digitized scene data may be stored in, for example, the frame buffer 330 in the set top box 140 or other processing device. Based on the pan/tilt/zoom command(s), a subset of the digitized scene image data is selected (625). In one embodiment, the webcam engine 410 processes the pan/tilt/zoom command(s) and selects the subset of the digitized scene image data based on the pan/tilt/zoom command(s).

Distortion compensation and compression is then performed (630) on the subset of the digitized scene image data. In one embodiment, the compression/correction engine 400 performs (630) the distortion compensation and compression of the subset of the digitized scene image data. The distortion-compensated and compressed subset is

then transmitted (635) to a selected destination such as display 120, to another device via Internet 150 or cable network 160, to printer 165, and/or to computer 170.

Figures 8A and 8B illustrate an example of another operation of embodiments of the invention. Assume the user sends a command 700 in order to capture an image of the object 710 and another command 705 to zoom the image of the object 710. A conventional webcam will require a physical pan movement to the left to capture the image of the object 705 and to capture a zoomed image of the object 705. Assume in this example that the digitized scene image data 485 of the scene in the vision field 200 was captured in the manner described above. The webcam engine 410 receives the pan left command 700 and accordingly selects an area 715 that contains an image of the object 710 in the digitized scene image data 485. The compression/correction engine 400 can perform distortion compensation to reverse the distortion caused by the wide angle lens 210 on the captured image of object 710. Typically, this compensation is performed by changing the curved surface of an image into a straight surface.

Also, as shown in Figure 8C, in response to the zoom command 705, the webcam engine 410 can enlarge an image of the selected area 715 in, for example, the frame buffer

330. The compression/correction engine 400 can then compress the image of selected area 715 and transmit the compressed image to a destination such as the display 120 or other suitable devices.

Reference is now made to Figures 8A and 9 to describe another function according to an embodiment of the invention. Assume the user sends a command 700 in order to capture an image of the object 710 and another command 740 to zoom away from the object 710. The webcam engine 410 receives the pan left command 700 and accordingly selects an area 750 that contains an image of the object 710 in the digitized scene image data 485. However, since the webcam engine 410 also received the zoom away command 740, the selected area 750 will be larger in size and cover a greater selected area portion in the digitized scene image area 485 than the selected area 715 in Figure 8B.

Figure 10 is a flowchart diagram of a method 800 to perform a zooming function according to an embodiment of the invention. A user first sends (805) a zoom command indicating whether to zoom in or away from an object to be captured in an image by a webcam. A scene in the field of vision of the lens of the webcam is then captured (810). The captured scene in the vision field is then stored (815) as scene image data in an image collection array. The

scene image data in the image collection array is then processed and stored (820) as a digitized scene image data. Based on the zoom command, a subset of the digitized scene image data is selected (825).

Processing of the subset of the digitized scene image data is then performed (827) based on the zoom command. For example, if the zoom command is for zooming the image of the captured object, then the subset of the digitized scene image data is enlarged. As another example, if the zoom command is for zooming away from the captured object, then the selected subset will cover a greater area in the digitized scene image data.

Distortion compensation and compression are then performed (830) on the subset of the digitized scene image data. The distortion-compensated and compressed subset is then transmitted (835) to a selected destination such as display 120, to another device via Internet 150 or cable network 160, to printer 165, and/or to computer 170.

Figure 11 is another diagram shown to further assist in describing an operation of an embodiment of the invention. A scene 900 falls within the vision field 905 of a wide angle lens 910 of a camera 915. The captured scene is digitized and processed into a digitized scene data 920. A subset 925 of the digitized scene data 920 is

selected based on a pan, tilt, and/or zoom command(s) that can be transmitted from an input device by the user. The selected subset 925 is then skew corrected (e.g., distortion compensated) into scene data 930 that can be transmitted to a destination. The scene data 930 is also typically compressed in order to optimize the data transmission across a network.

Other variations and modifications of the above-described embodiments and methods are possible in light of the foregoing teaching. For example, webcam 100 may comprise a processor and perform the selection of the subset of the digitized scene image data and the distortion compensation and compression of the subset instead of STB 140. As another example, the webcam 100 can send the digitized scene image output to a processing device, such as a personal computer instead of the STB 140, and the processing device can select the subset of the digitized scene image data and perform the distortion compensation and compression of the subset.

As another example, the webcam 100 can instead send the digitized scene image output to an optional companion box device 175 (Figure 1) instead of sending the digitized scene image output to the set top box 140. The companion box 175 may include, for example, the functionality of an

Interactive Companion Box, as described in U.S. Patent Application No. ___/_______, filed on March 22, 2001, entitled "Interactive Companion Set Top Box," by inventors Ted M. Tsuchida and James A. Billmaier, the disclosure of which is hereby incorporated by reference. Functions of the Interactive Companion Box may include Internet access, Video-on-Demand, an electronic programming guide, videoconferencing, and/or other functions.

As another example, the sample stage 245 in Figure 1 may instead perform the selection of the image subset to be compressed and compensated for distortion, instead of the webcam engine 410.

Further, at least some of the components of this invention may be implemented by using a programmed general purpose digital computer, by using application specific integrated circuits or field programmable gate arrays, or by using a network of interconnected components and circuits. Connections may be wired, wireless, by modem, and the like.

The above description of illustrated embodiments of the invention, including what is described in the Abstract, is not intended to be exhaustive or to limit the invention to the precise forms disclosed. While specific embodiments of, and examples for, the invention are described herein

for illustrative purposes, various equivalent modifications are possible within the scope of the invention, as those skilled in the relevant art will recognize.

These modifications can be made to the invention in light of the above detailed description. The terms used in the following claims should not be construed to limit the invention to the specific embodiments disclosed in the specification and the claims. Rather, the scope of the invention is to be determined entirely by the following claims, which are to be construed in accordance with established doctrines of claim interpretation.